UMassAmherst

Joint Alignment

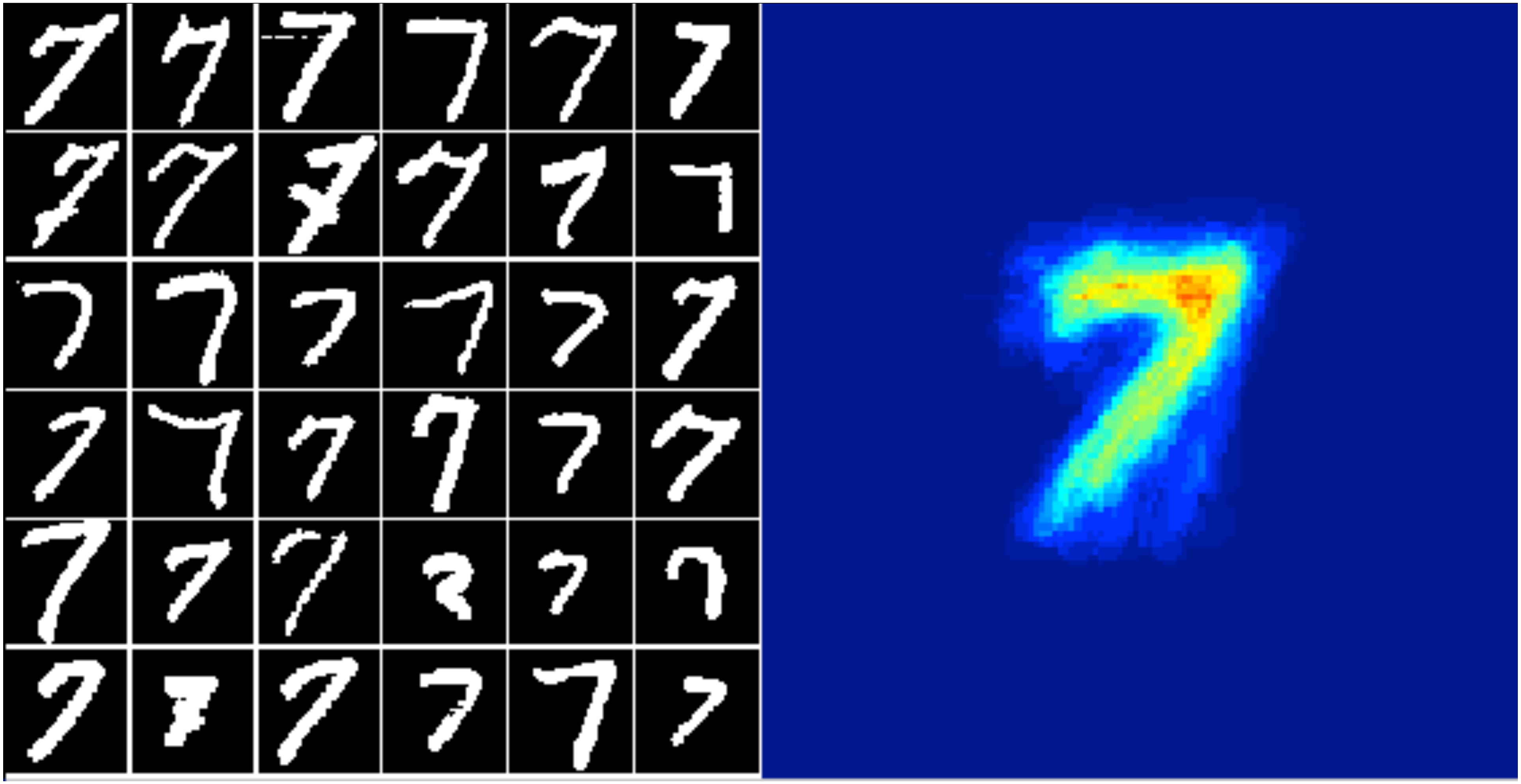
Including work with

Vidit Jain, Andras Ferencz, Gary Huang, Lilla Zollei, Sandy Wells

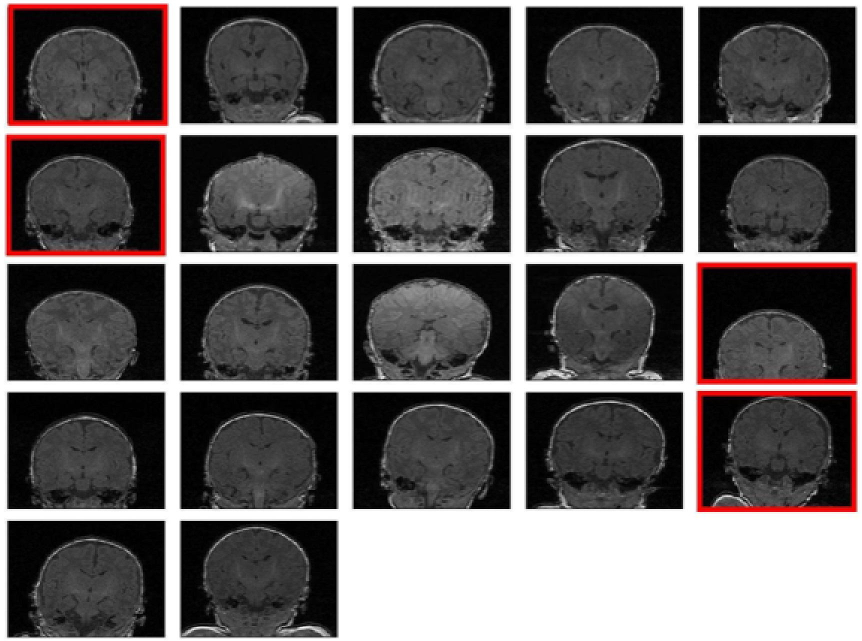
Examples of Joint Alignment

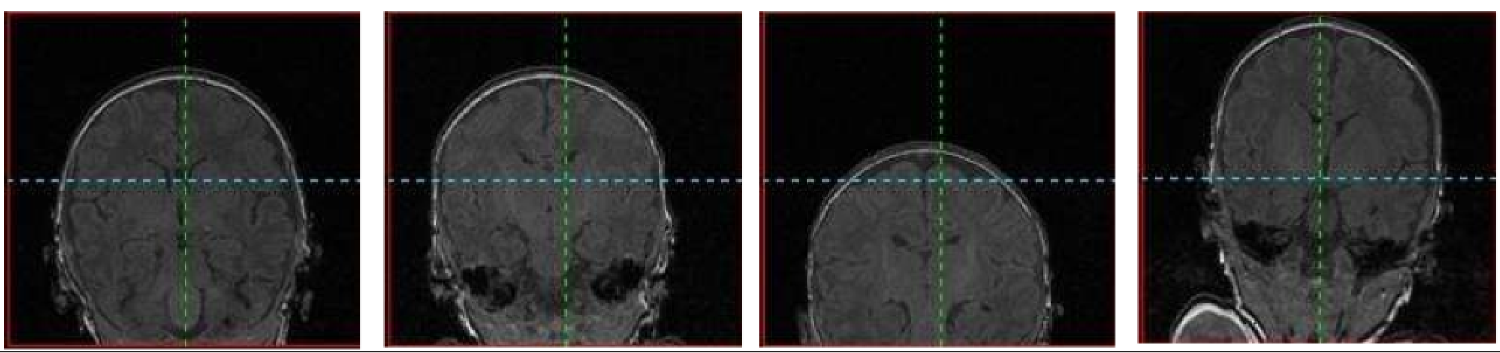
- Aligning handwritten digits
  - Improves recognition
  - Allows recognition from a single example
- Aligning grayscale images and grayscale volumes
  - magnetic resonance images
- Aligning complex images such as faces
  - Improves recognition
  - Building a hierarchy of models, from coarse to fine

Congealing *(CVPR 2000, PAMI 2006)*

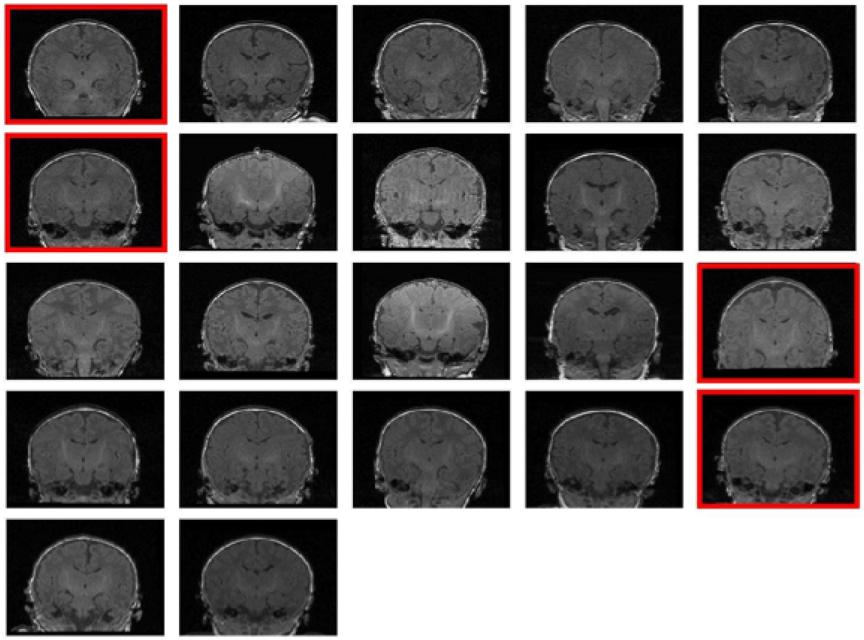


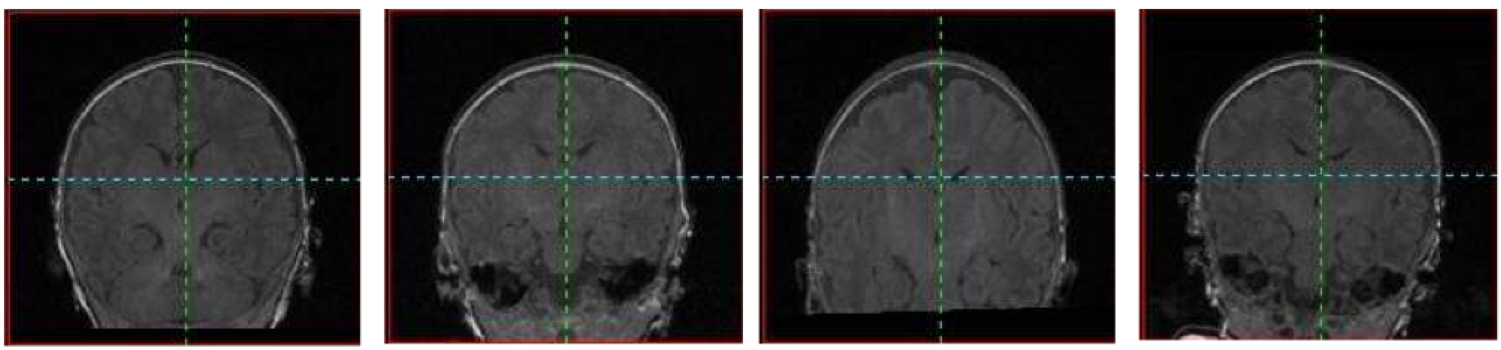
Congealing Gray Brain Volumes *(ICCV 2005 Workshop)*





Aligned Volumes





Why joint alignment?

- Can be easier than aligning two images!
  - Natural smoothing effect.
- Produces natural notion of "center".
  - Traditional medical atlas: one individual
  - Compares anatomy to many individuals that have been jointly registered
- Automatically produce an alignment machine (an "image funnel") from a set of images.
  - Unsupervised model building!
- Produce "sharper" models.

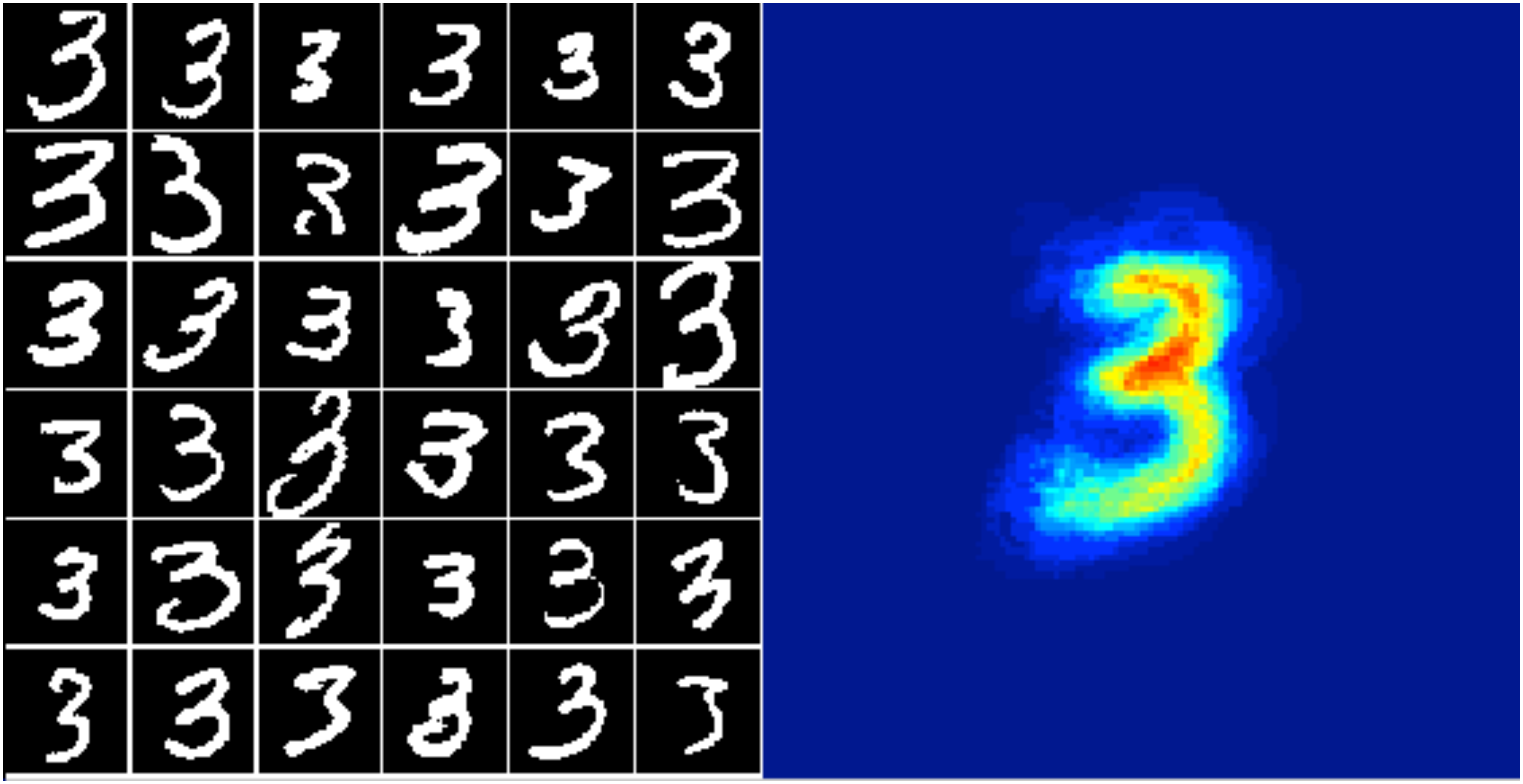
Congealing

- Process of joint alignment of sets of arrays (samples of continuous fields).
- 3 ingredients
  - A set of arrays in some class
  - A parameterized family of *continuous* transformations
  - A criterion of joint alignment

Congealing Binary Digits

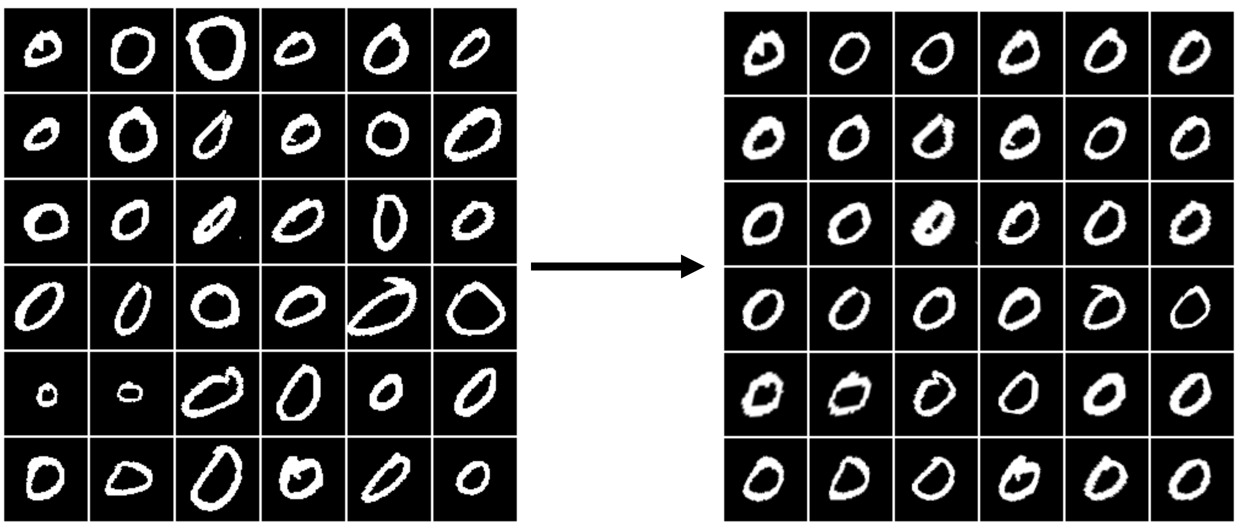
- 3 ingredients
  - A set of arrays in some class:
    - Binary images
  - A parameterized family of *continuous* transformations:
    - Affine transforms
  - A criterion of joint alignment:
    - Entropy minimization

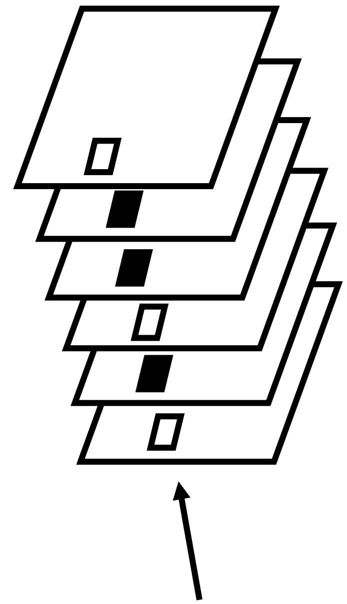
Congealing



Criterion of Joint Alignment

- Minimize sum of pixel stack entropies by transforming each image. “Joint Gradient Descent”





Entropy

Entropy of a **discrete random variable** X that takes values in $\mathcal{X}$:

$$H(X) = -\sum_{x \in \mathcal{X}} P(x) \log P(x) \quad (1)$$

$$= -E[\log P(X)]. \quad (2)$$

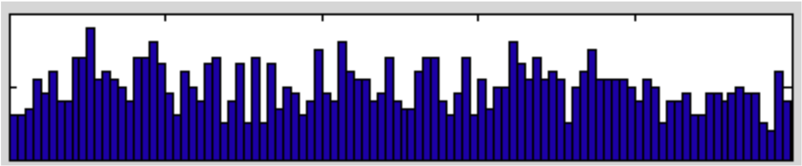
Differential entropy of a **continuous real random variable** X:

$$h(X) = -\int_{-\infty}^{\infty} p(x) \log p(x) \quad (3)$$

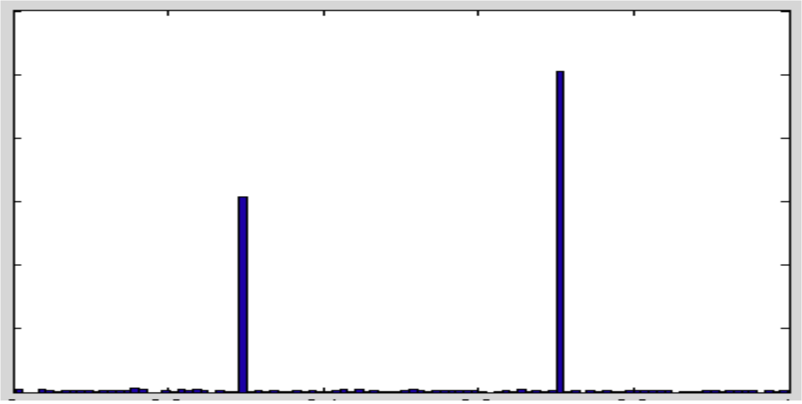
$$= -E[\log p(X)]. \quad (4)$$

Entropy of probability distributions

Histogram of samples from a high entropy distribution.



Histogram of samples from a low entropy distribution.



Entropy as a measure of dispersion

- Low entropy
  - High average log likelihood under "true" distribution.
  - A small number of highly likely values
- High entropy
  - a large number of relatively uncommon values.

- Important for gray scale images:
  - Multi-modal distribution can have low entropy!
    - Even if the modes are far apart.
  - Variance does not have this property!

Empirical entropy

- Empirical entropy is the estimate of the entropy of a random variable derived from a sample.

  - Given: A sample of a random variable X.
  - To estimate entropy of X:
    - Estimate probability distribution of X from the sample (density estimation).
    - Compute the entropy of the density estimate.

Empirical entropy

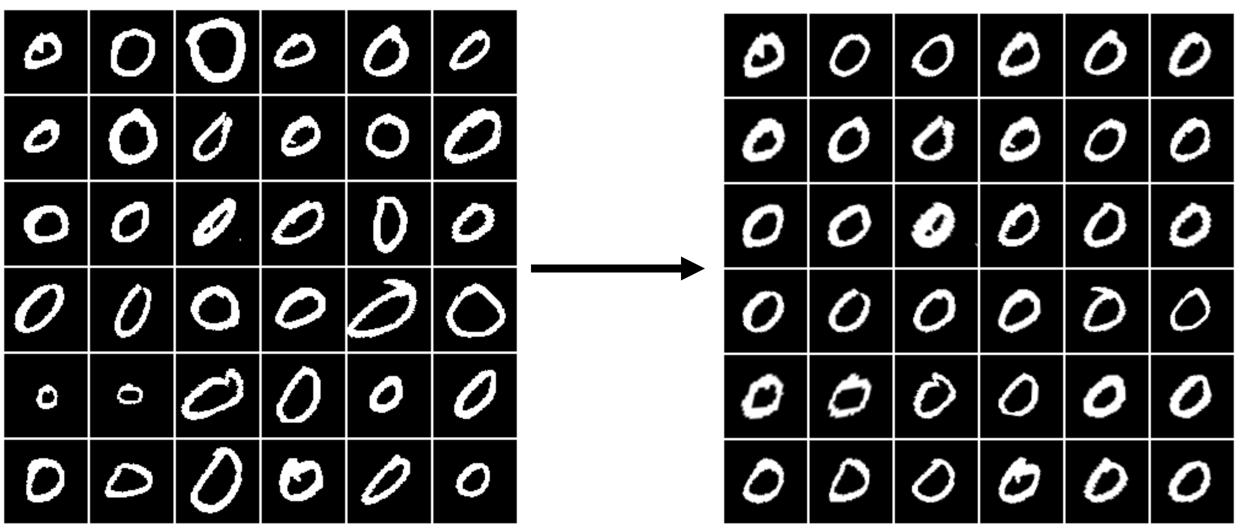
- Empirical entropy is the estimate of the entropy of a random variable derived from a sample.

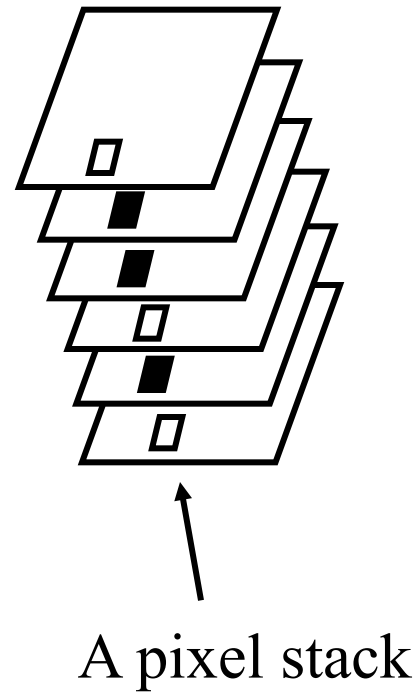
  - Given: A sample of a random variable X.
  - To estimate entropy of X:
    - Estimate probability distribution of X from the sample (density estimation).
    - Compute the entropy of the density estimate.

There are very fast methods of entropy estimation that do not require the intermediate estimation of a density!

Criterion of Joint Alignment

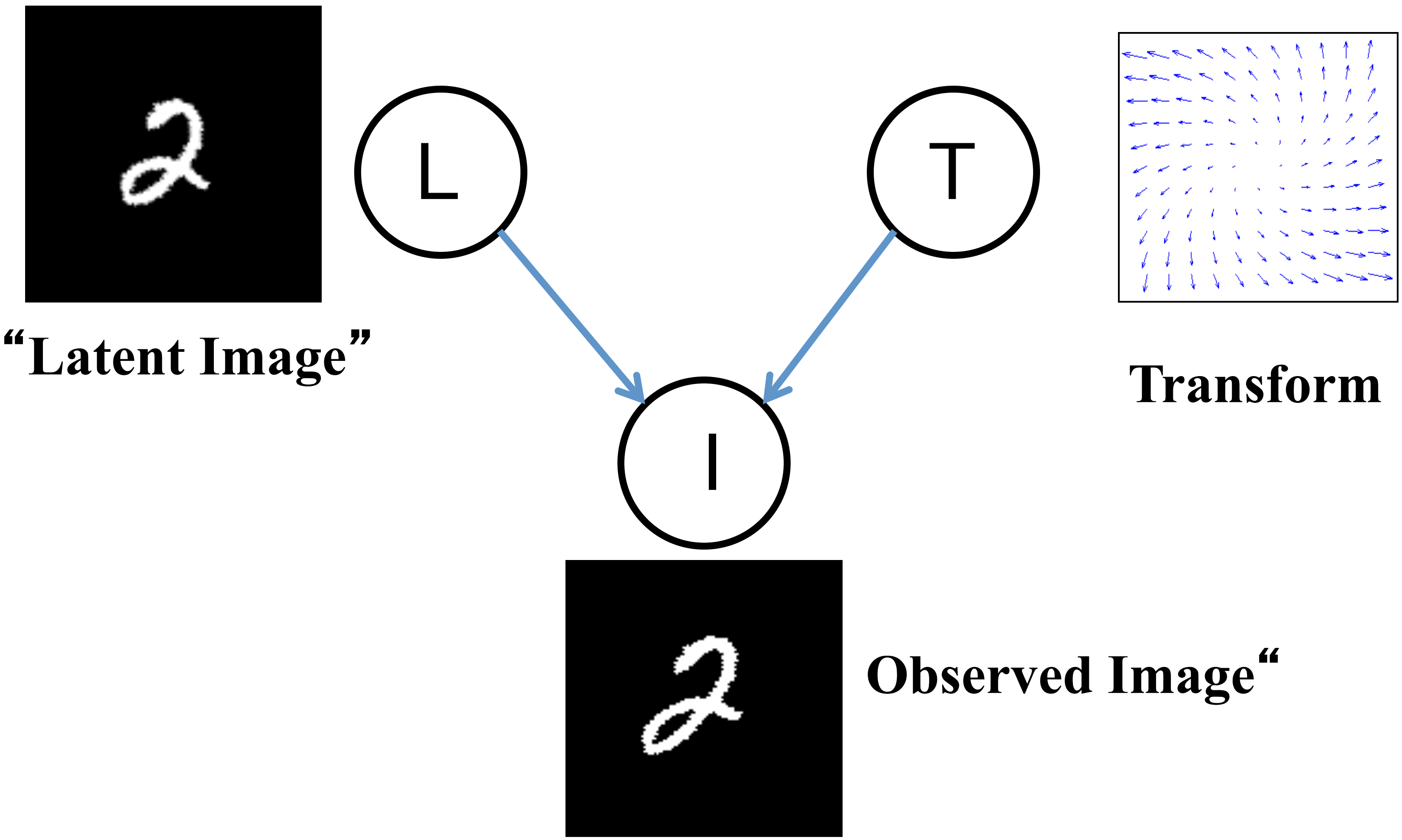
- Minimize sum of pixel stack entropies by transforming each image.



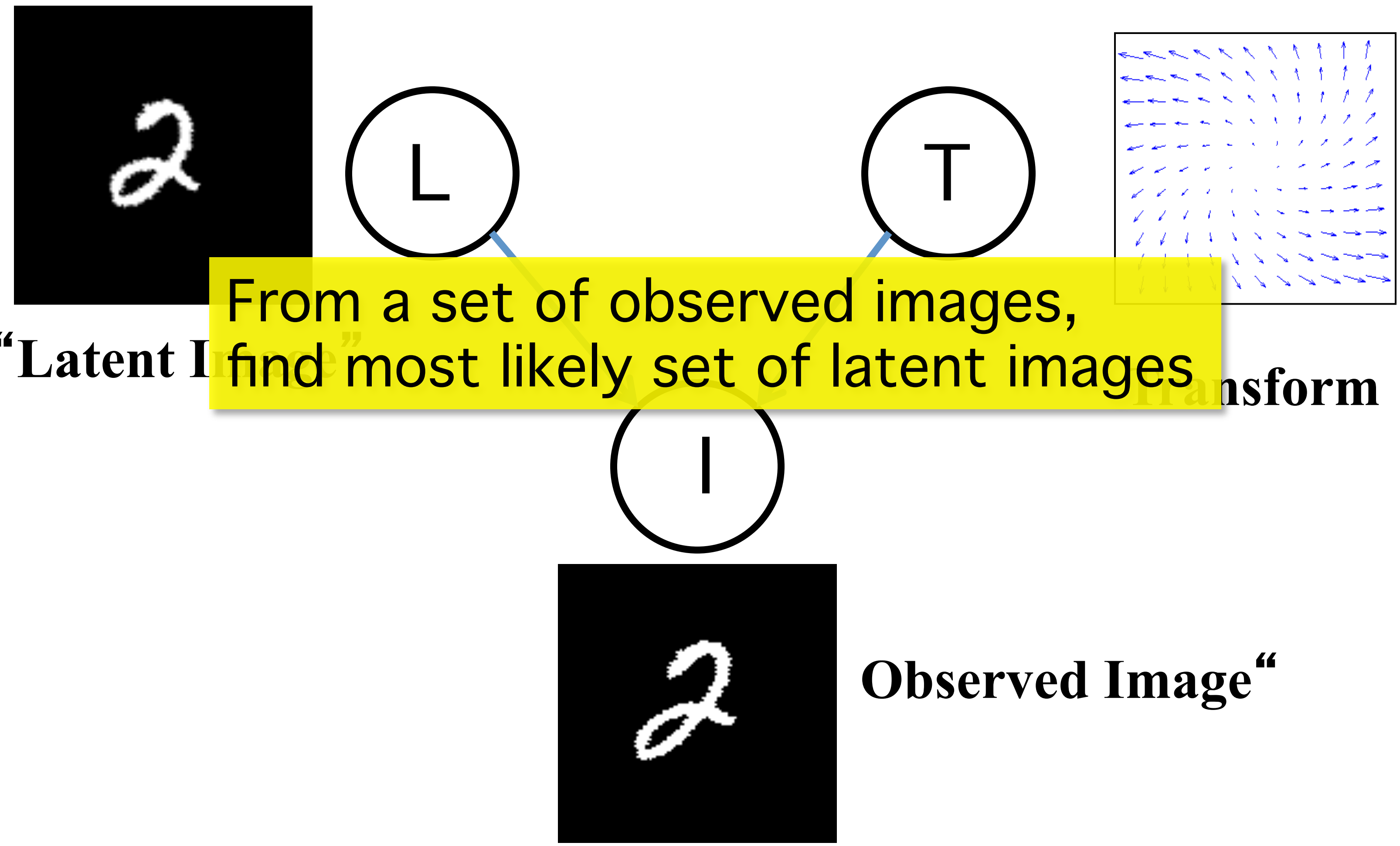


Note: Mutual Information doesn't make sense here.

Congealing as Inference



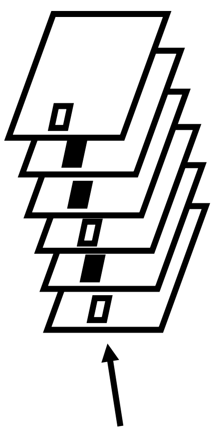
Congealing as Inference



From a set of observed images, find most likely set of latent images

Min entropy = Max non-parametric likelihood

$$
\begin{aligned}
\arg\max_{\mathbf{T}\in\mathcal{T}} P(\mathbf{T}|\mathbf{I}) &= \arg\max_{\mathbf{T}\in\mathcal{T}} P(\mathbf{I},\mathbf{T}) && (1)\\
&\approx \arg\max_{\mathbf{T}\in\mathcal{T}} P(\mathcal{L}(\mathbf{I},\mathbf{T})) && (2)\\
&= \arg\max_{\mathbf{T}\in\mathcal{T}} \prod_{x,y}\prod_{i=1}^{N} P_{x,y}(L_i(x,y)) && (3)\\
&= \arg\max_{\mathbf{T}\in\mathcal{T}} \sum_{x,y}\sum_{i=1}^{N} \log P_{x,y}(L_i(x,y)) && (4)\\
&\approx \arg\min_{\mathbf{T}\in\mathcal{T}} -\sum_{x,y}\sum_{i=1}^{N} \log \hat{P}_{x,y}(L_i(x,y)) && (5)\\
&= \arg\min_{\mathbf{T}\in\mathcal{T}} \sum_{x,y} \hat{H}(X,Y) && (6)
\end{aligned}
$$



A pixel stack

The Independent Pixel Assumption

- Model assumes independent pixels
- A poor generative model:
  - True image probabilities don't match model probabilities.
  - Reason: heavy dependence of neighboring pixels.
- However! This model is great for alignment and separation of causes!
  - Why?
  - Relative probabilities of "better aligned" and "worse aligned" are usually correct.

Summary so far…

- Congealing aligns a set of images
- It does this by trying to make each column of pixels (a pixel stack) have low disorder (entropy)
- It assumes that the distribution of latent images have independent pixels.

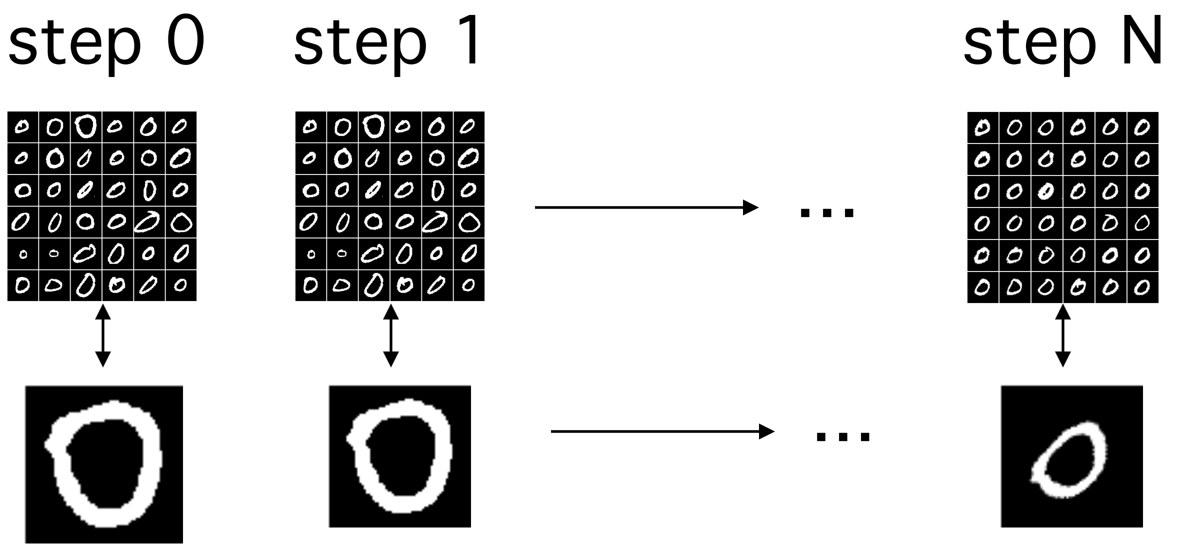
Summary so far…

- Congealing aligns a set of images
- It does this by trying to make each column of pixels (a pixel stack) have low disorder (entropy)
- It assumes that the distribution of latent images have independent pixels.

- Next question: what if we want to align one new image to the set of images we have already aligned?

How do we align a new image?

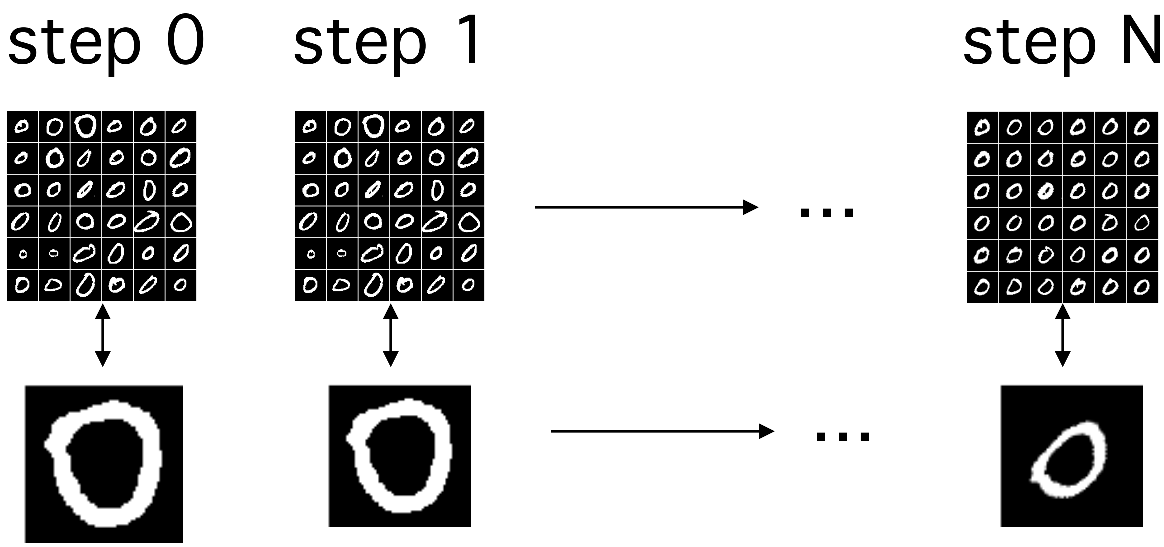
Sequence of successively "sharper" models



Take one gradient step with respect to each model.

How do we align a new image?

Sequence of successively "sharper" models



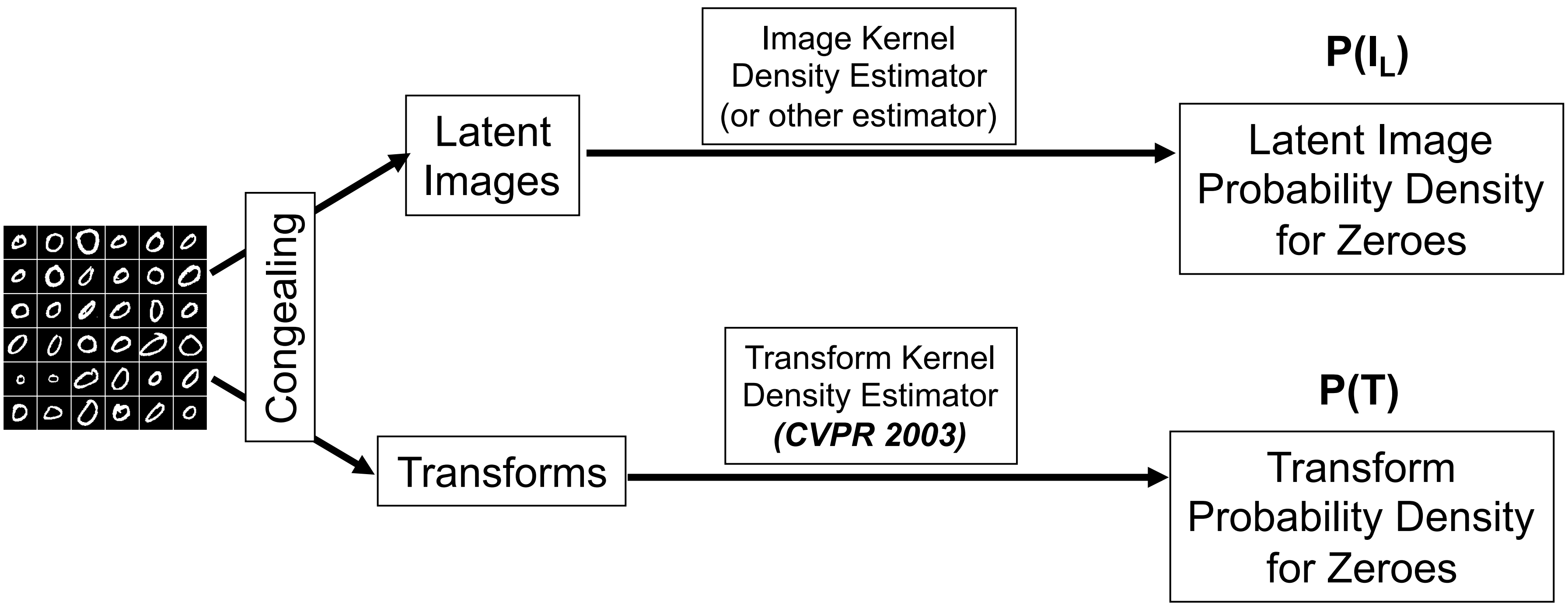
Funneling

- A funnel is an image alignment machine.
- It is a side-effect of the congealing process.
- Congealing any set of images produces a funnel which can be used align subsequent images

- NO TRAINING DATA ARE REQUIRED!!!

Applications…

Learning from one example (CVPR 2000)



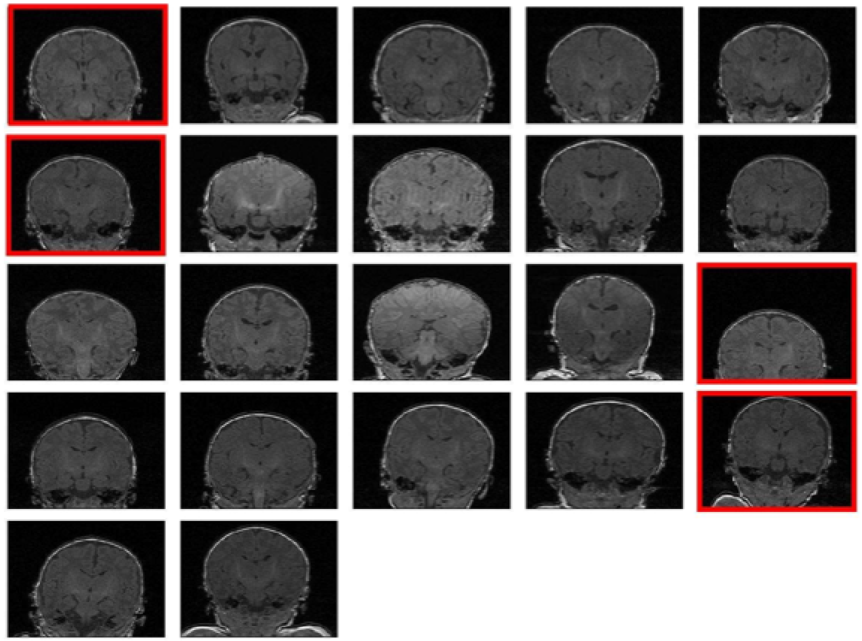
Application: Alignment of 3D Magnetic Resonance Volumes

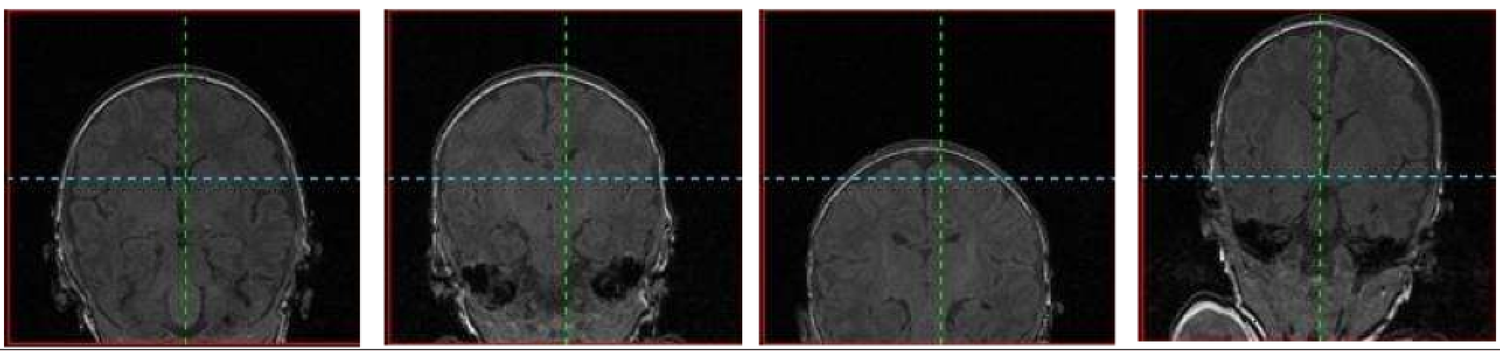
Lilla Zollei, Sandy Wells, Eric Grimson

Congealing MR Volumes: Joint Registration

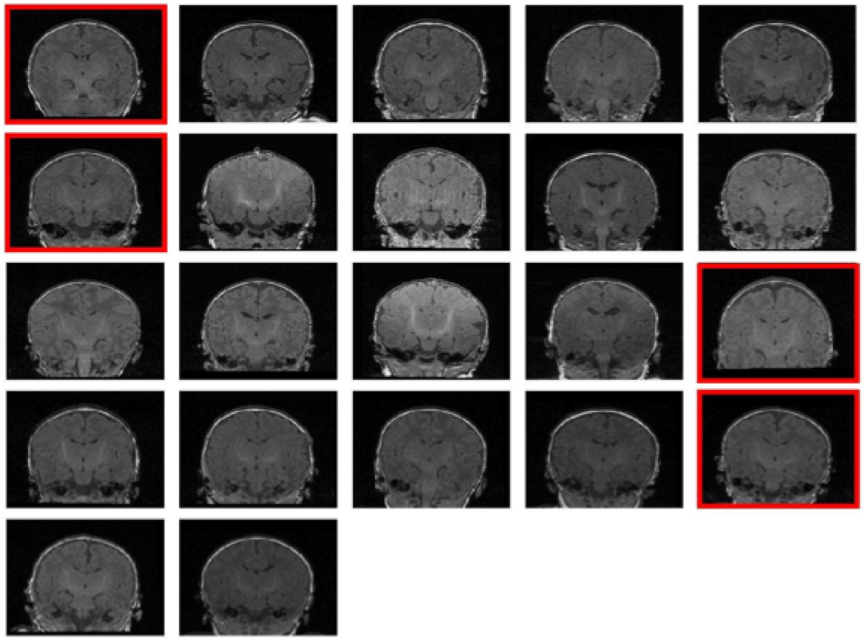
- 3 ingredients
  - A set of arrays in some class:
    - Gray-scale MR volumes
  - A parameterized family of *continuous* transformations:
    - 3-D affine transforms
  - A criterion of joint alignment:
    - Grayscale entropy minimization
- Purposes:
  - Pooling data for functional MRI studies
  - Aligning subjects to a common **unbiased** reference frame for comparison
  - Building general purpose statistical anatomical atlases

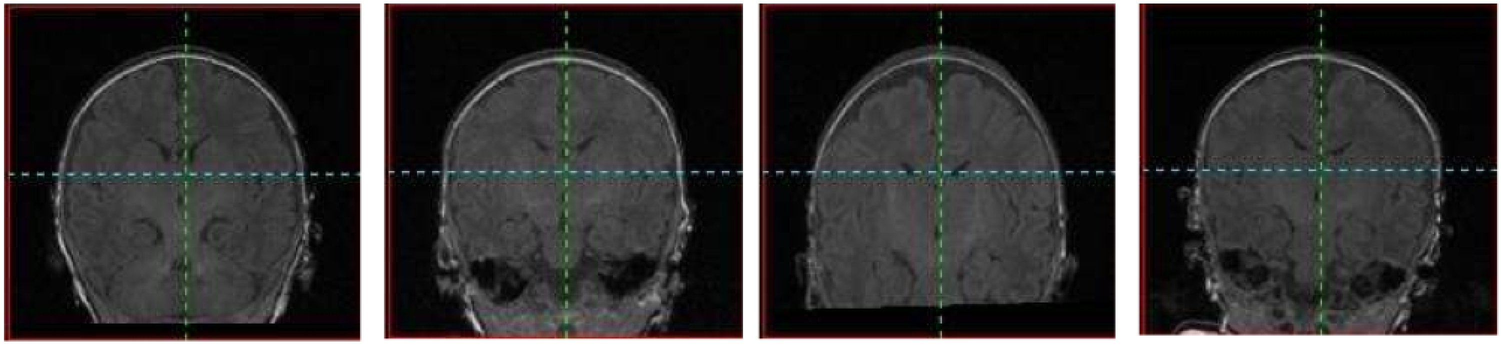
Congealing Gray Brain Volumes *(ICCV 2005 Workshop)*



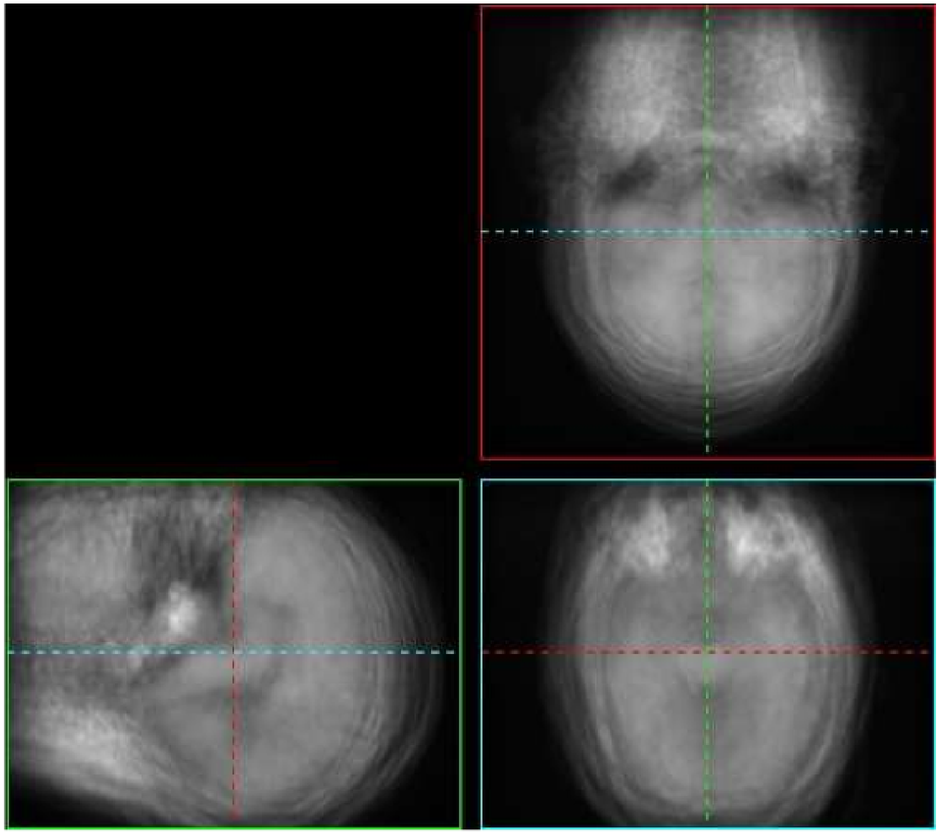


Aligned Volumes

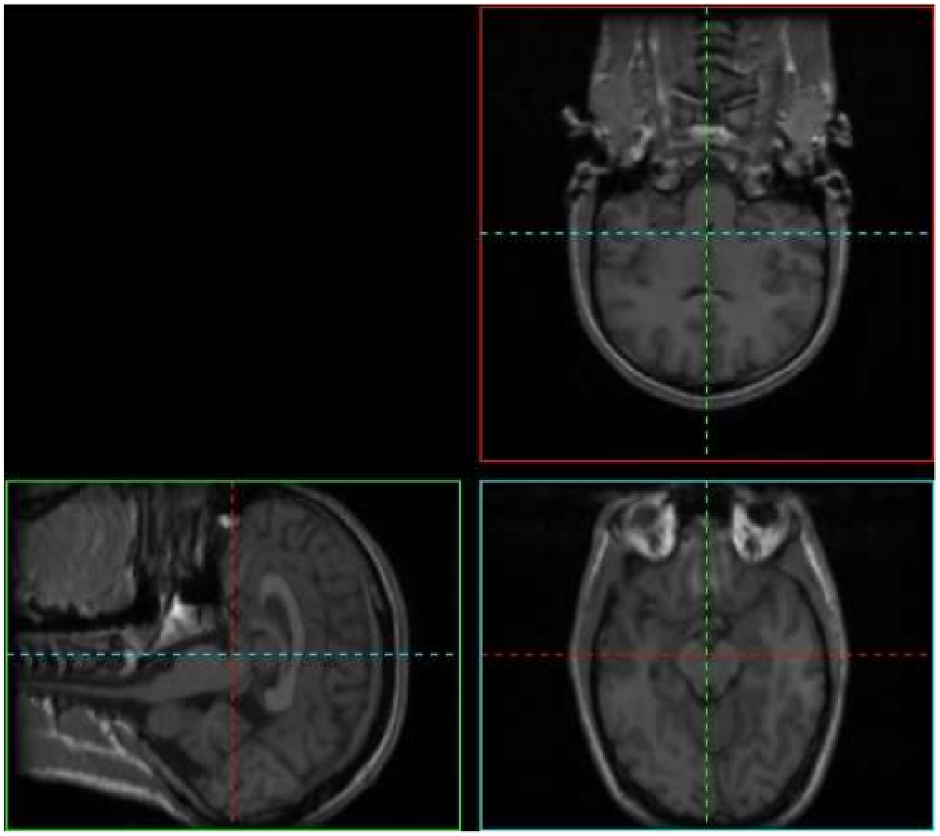




Validation: Synthetic Data

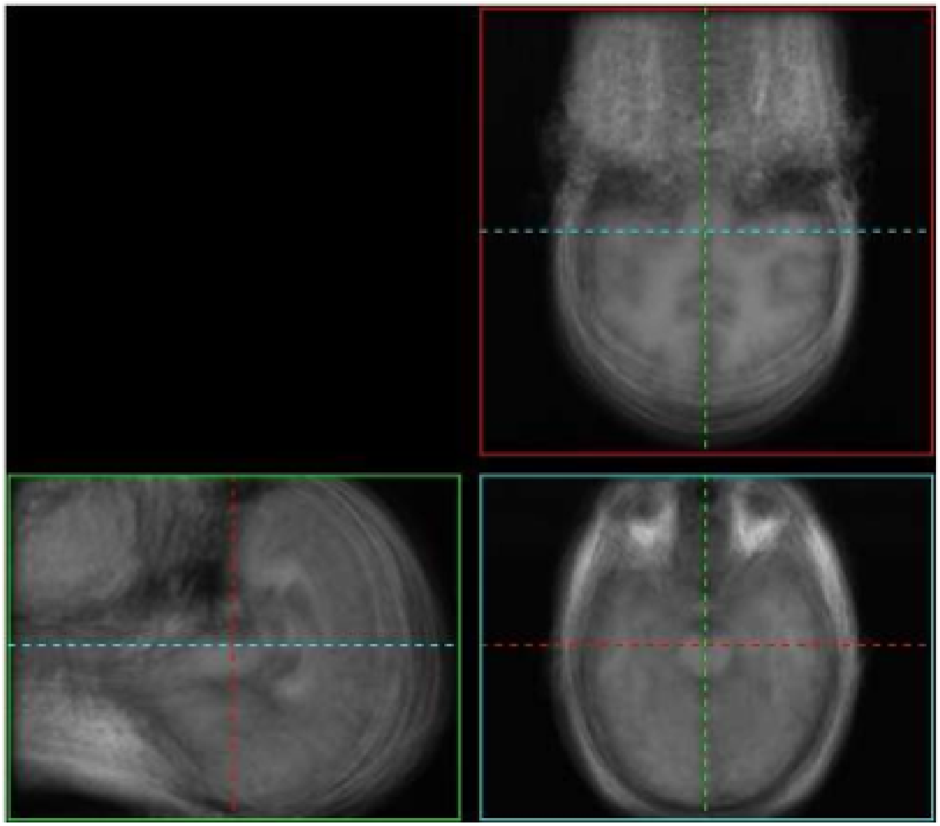


Unaligned input data sets

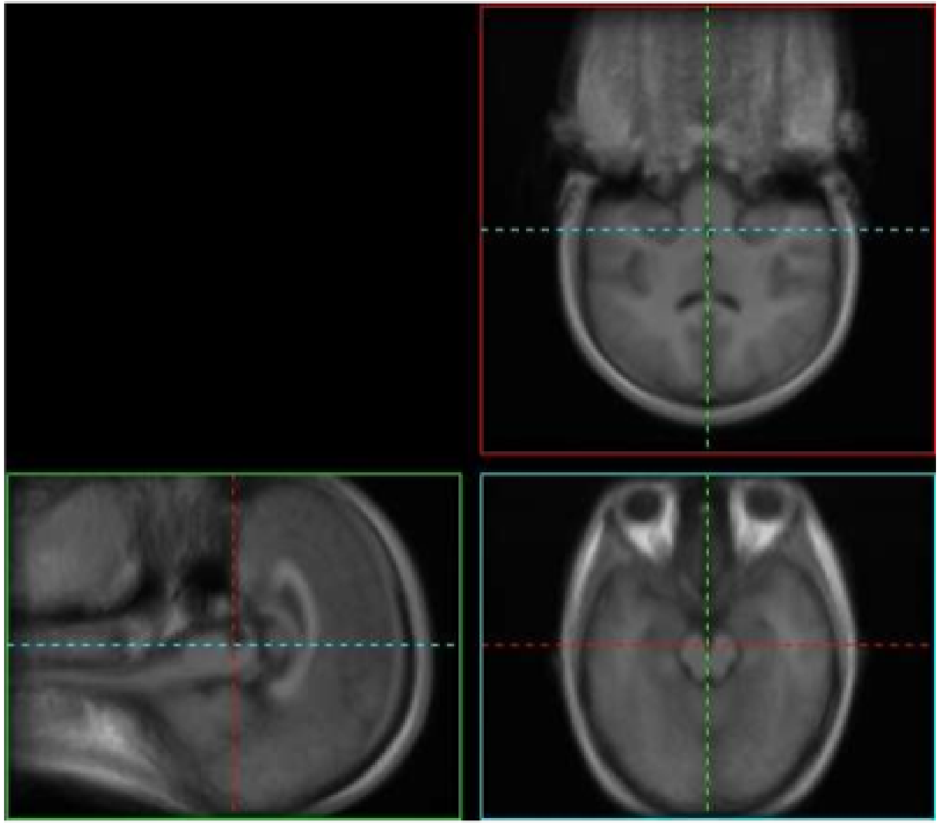


Aligned input data sets

Real Data



Unaligned input data sets



Aligned input data sets

Data set: 28 T1-weighted MRI; [256x256x124] with (.9375, .9375, 1.5) mm^3 voxels

Experiment: 2 levels; 12-param. affine; N = 2500; iter = 150; time = **1209 sec!!**

MR Congealing Challenges

- Big data
  - 8 million voxels per volume
  - 100 volumes
  - 12 transform parameters (3D affine)
  - 20 iterations
- Techniques:
  - Stochastic sampling
  - Multi-resolution techniques

Next Application:
Alignment of Faces for Improved Recognition

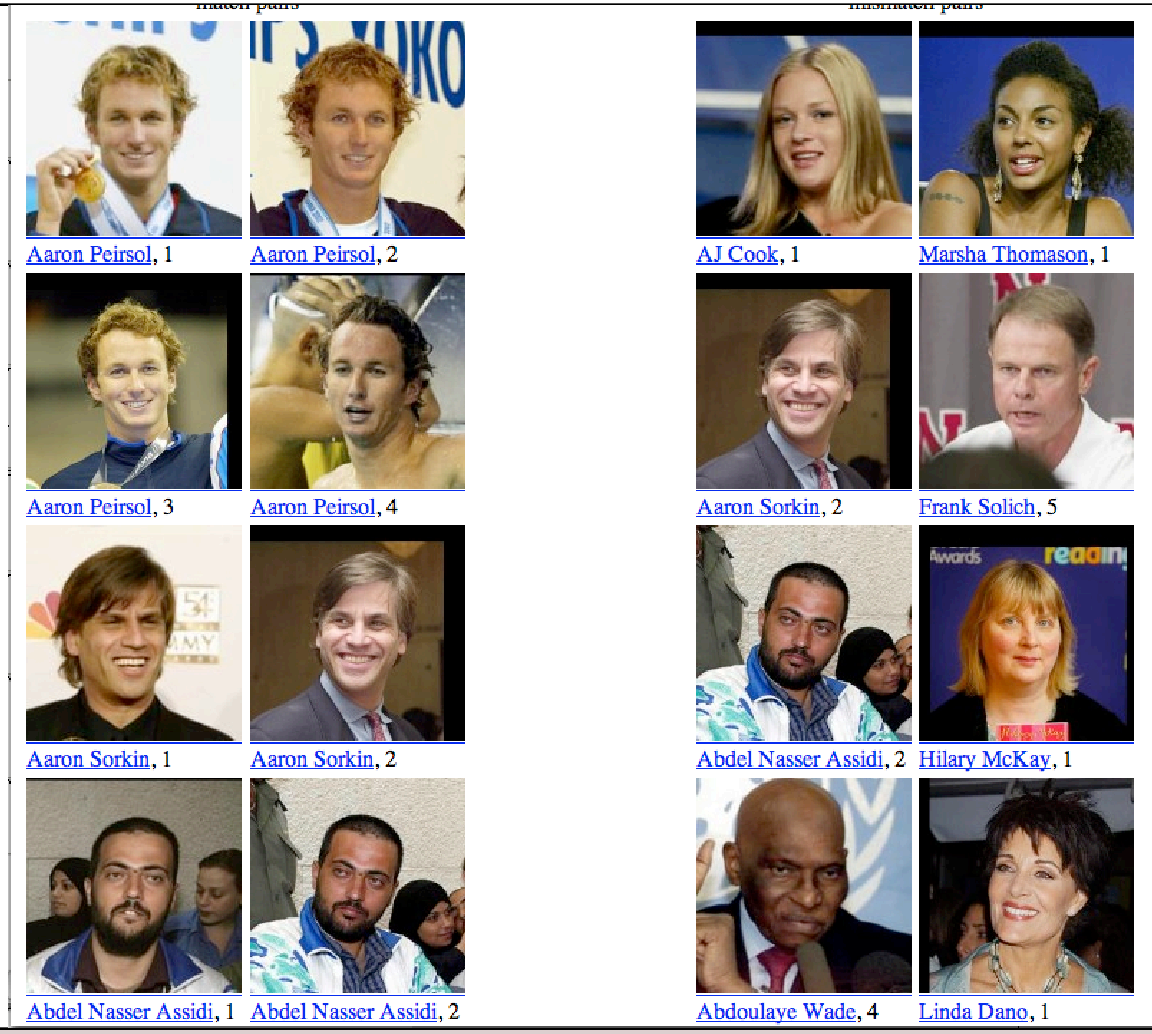
joint work with Gary Huang

Labeled Faces in the Wild

http://vis-www.cs.umass.edu/lfw/

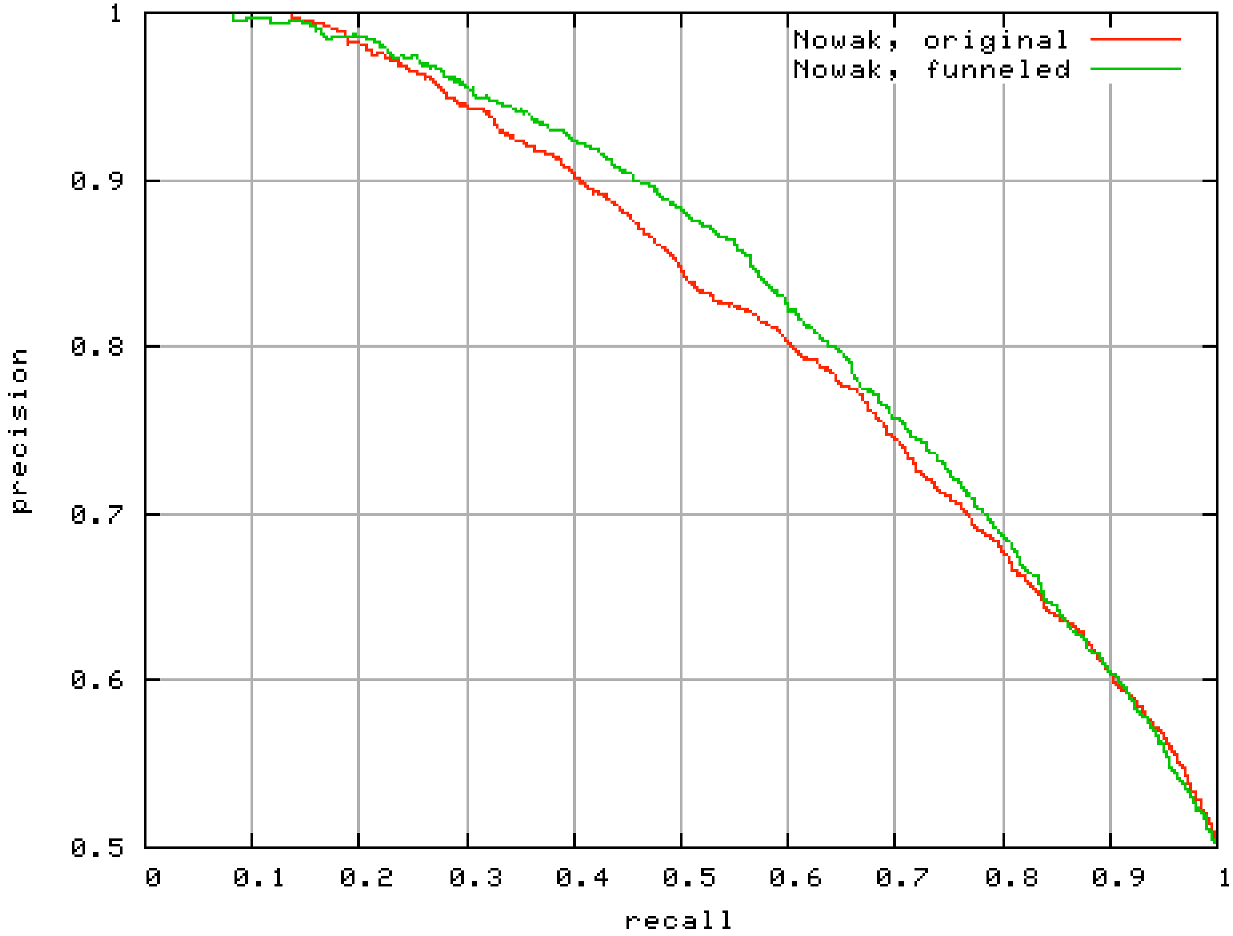
Labeled Faces in the Wild: Face Verification

“same”



“different”

Face verification with and without alignment



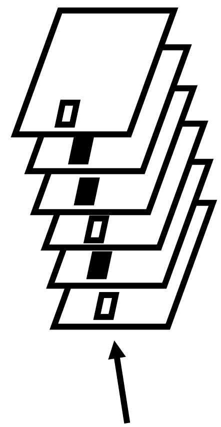
Traditional Face Alignment

- Traditional Face Alignment algorithm:
  - Develop "part detectors" for eyes, nose, mouth, and other parts of the face.
    - Requires lots of hand-labeled data.
  - Find the parts for a new face.
  - Position those parts in canonical locations.

- Is it possible to design an alignment algorithm without first building part detectors?
  - An "unsupervised" alignment algorithm.
    - Unsupervised because no parts were labeled.

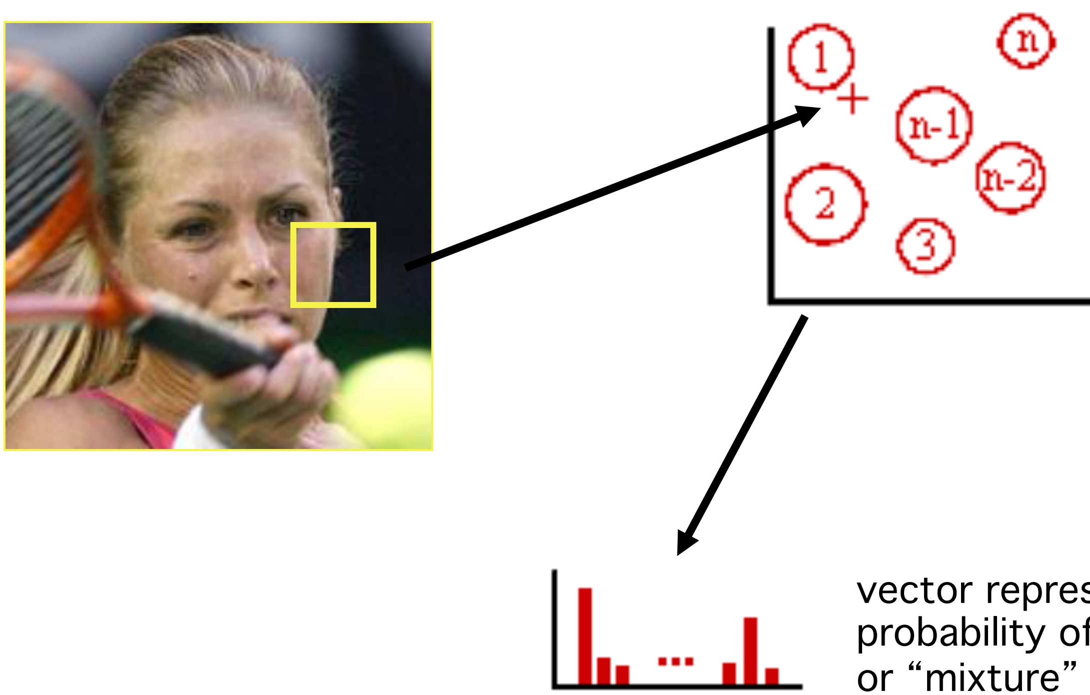
Congealing Faces

- Challenges:
  - High variability
  - Pixel values do not necessarily have low entropy when aligned
    - Lighting, hue may foil pixel-based method
- Use higher level-features that have greater invariance under lighting
  - SIFT (what else?)
- Problem with SIFT—high dimensionality
  - Can't estimate entropy of SIFT distribution from small number of examples.
  - Need to reduce dimensionality



A pixel stack

Congealing Complex Images (ICCV 2007)



SIFT clusters

vector representing
probability of each cluster,
or “mixture” of clusters

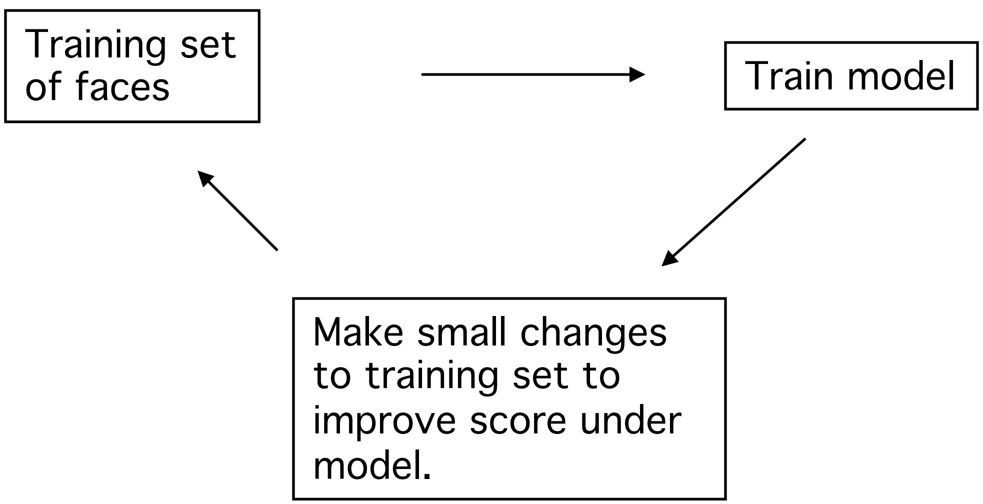
Convert face images to arrays of multinomials

- Start with data set of faces
- Compute SIFT at each pixel
- Cluster SIFT vectors (16 clusters)
- At each pixel, form posterior (multinomial) over clusters
- Distribution of pixel stack is mean of multinomial vectors
- Now, do congealing over these multinomial vectors

Face Congealing

Converting any Model into a Congealer

- Congealing as sequence of independent pixel models.
  - Why not use other models?
    - For example, PCA congealing?



Deep Congealing (NIPS 2012)

- Build a model of faces using Deep Belief Networks.
- Adjust each face to increase its likelihood under the Deep Belief model.
- (Retrain the Deep Belief model).
- Iterate until convergence

Deep Congealing (in submission)

- Build a model of faces using Deep Belief Networks.
- Adjust each face to increase its likelihood under the Deep Belief model.
- Retrain the Deep Belief model.
- Iterate until convergence

- Matches best alignment performance so far, but with no annotated parts!

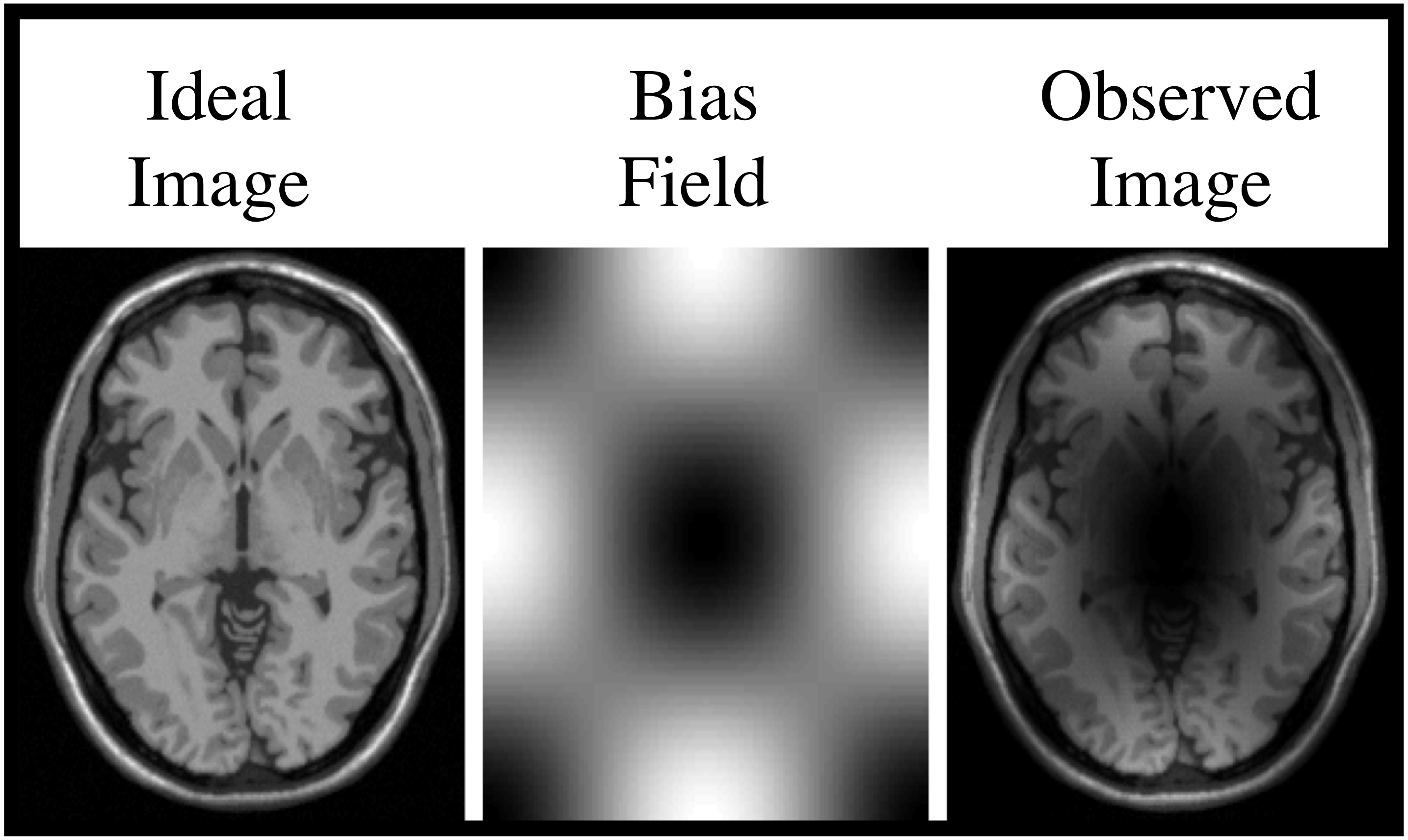
Deep Congealing

Summary of Face Congealing

- Fine alignment significantly increases recognition rates for most face recognition algorithms.
- Congealing can be done in different feature spaces
  - Must be able to estimate entropy of feature space from a few hundred examples at most
- Congealing can be done with respect to different models
  - Deep Congealing
- Nothing in the algorithm is specific to faces
  - Works just as well with frontal car images!

Last Application:
Bias removal in MRI

The Problem

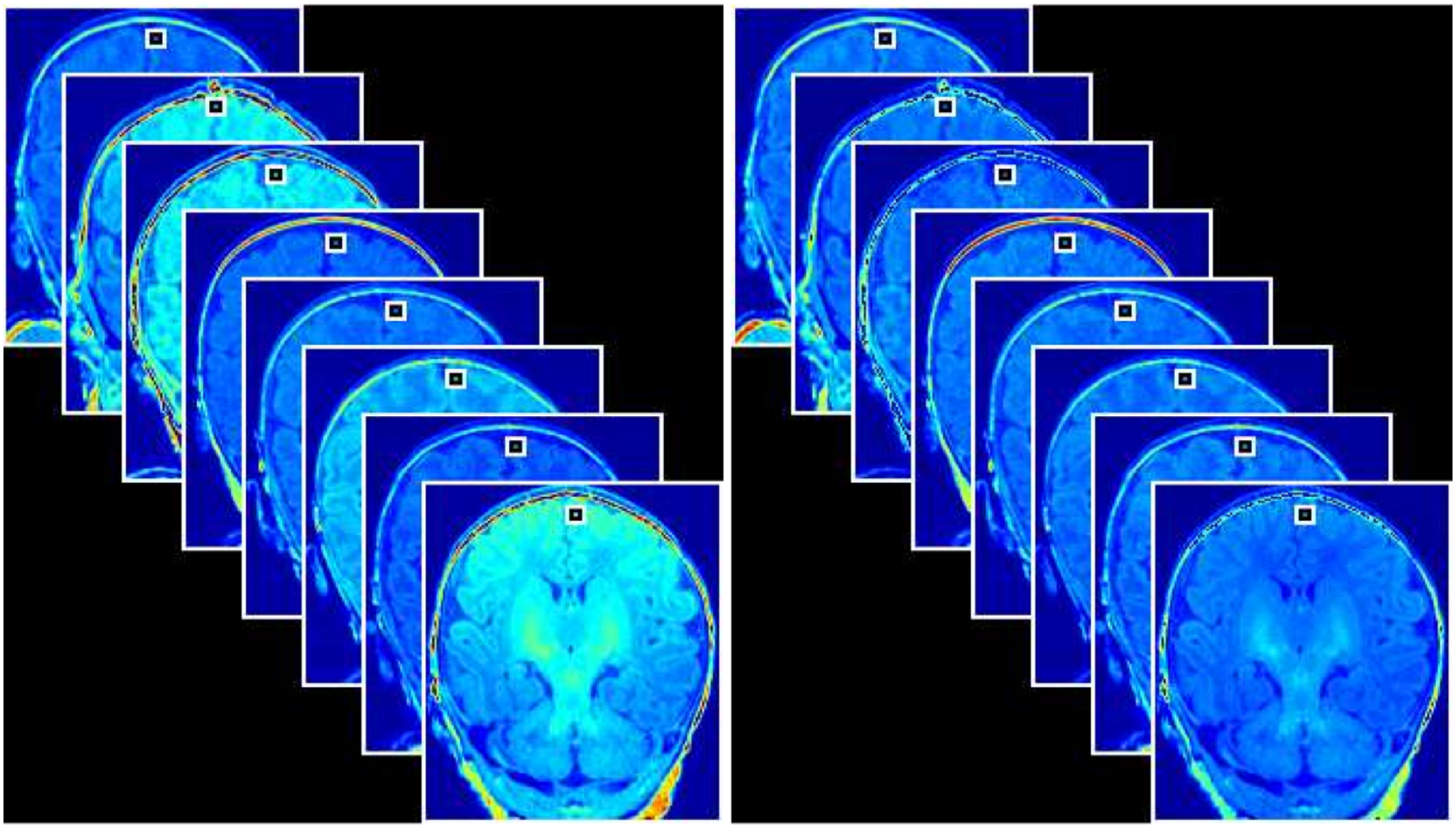


Bias fields have low spatial frequency content

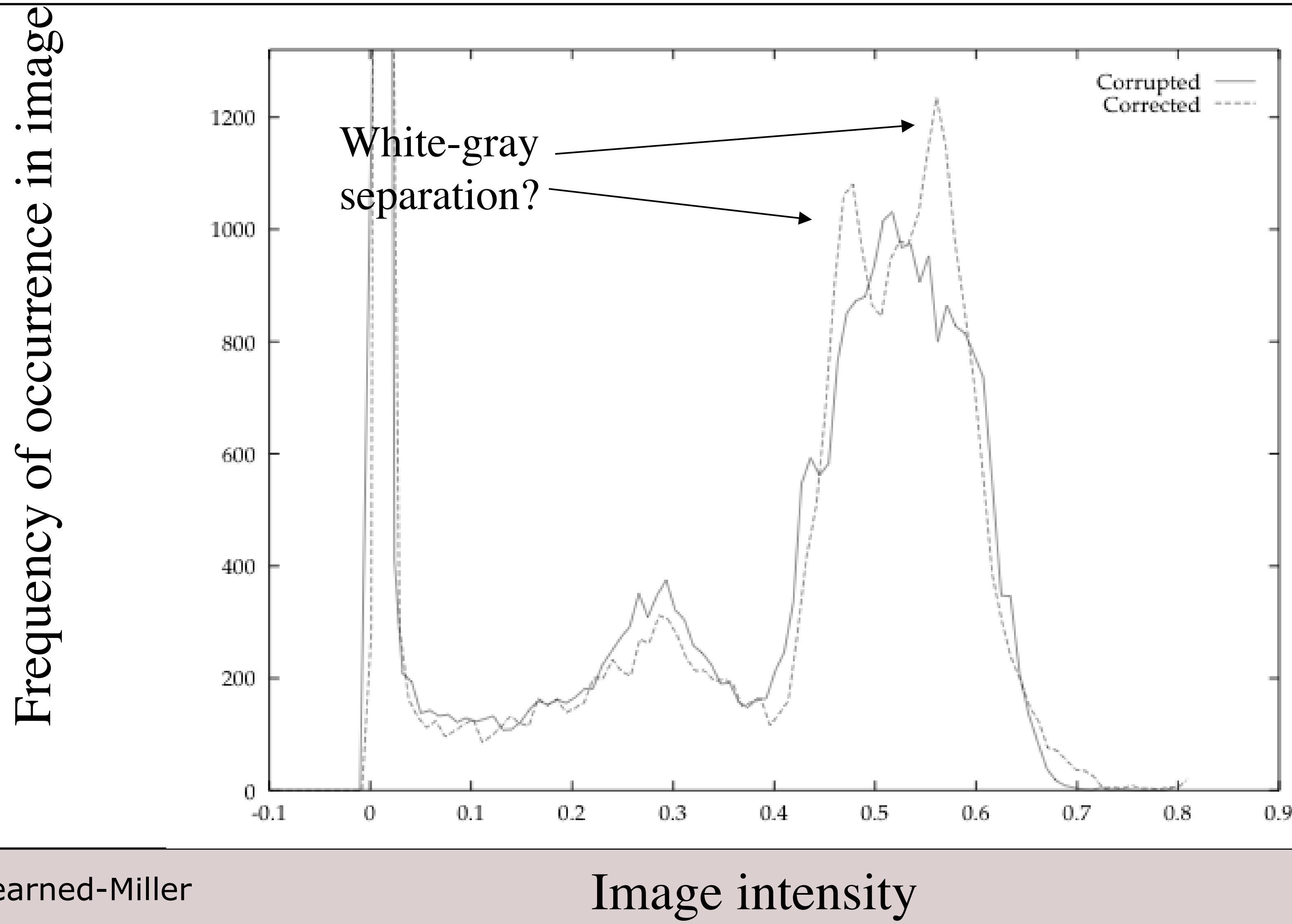
Bias Removal in MR as a Congealing Problem

- 3 ingredients
  - A set of arrays in some class:
    - MR Scans of Similar Anatomy (2D or 3D)
  - A parameterized family of *continuous* transformations:
    - Smooth brightness transformations
  - A criterion of joint alignment:
    - Entropy minimization

Congealing with brightness transforms

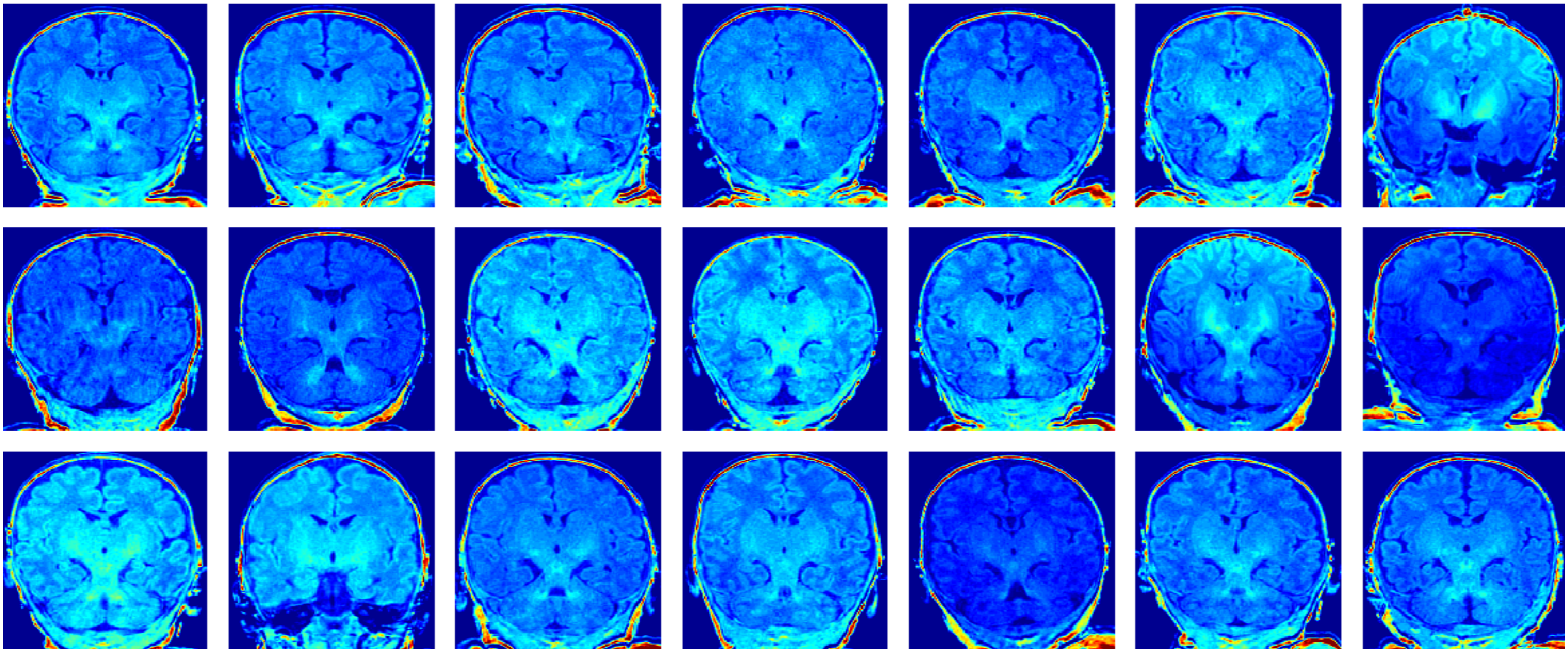


Grayscale Entropy Minimization



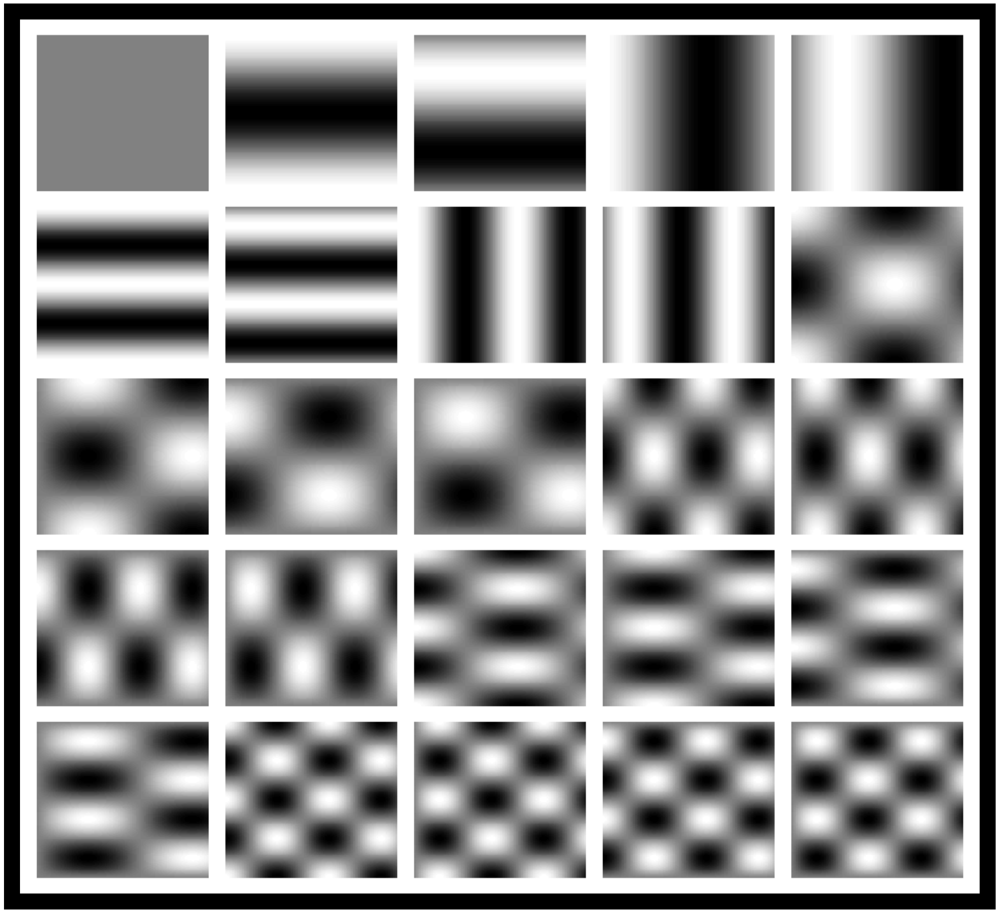
Some Infant Brains

(thanks to Inder, Warfield, Weisenfeld)



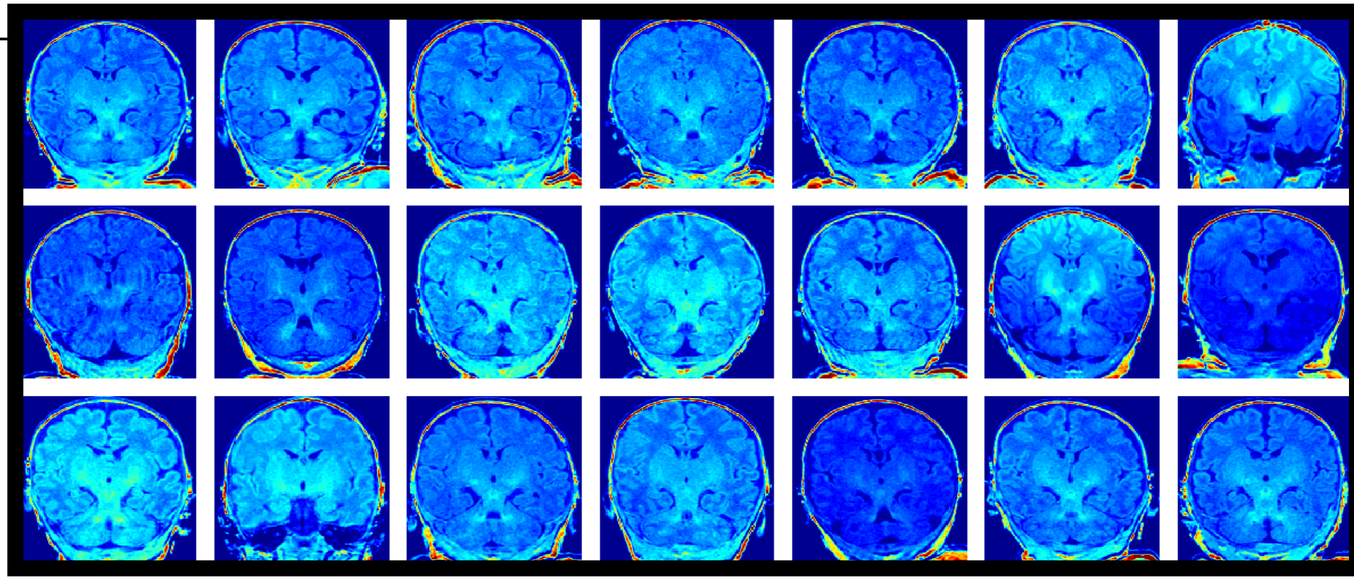
- Pretty well registered (not perfect)
- Pretty bad bias fields

Fourier Basis for Smooth Bias Fields

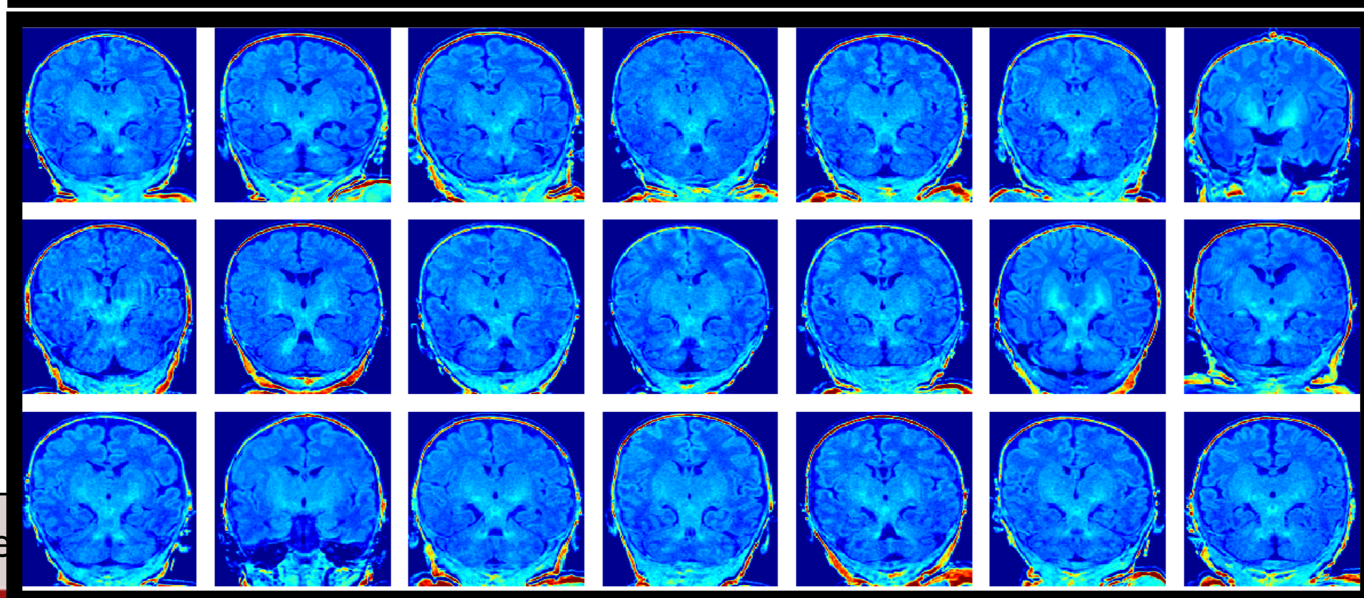


Results

Original Images



Bias Corrected Images



Assumptions

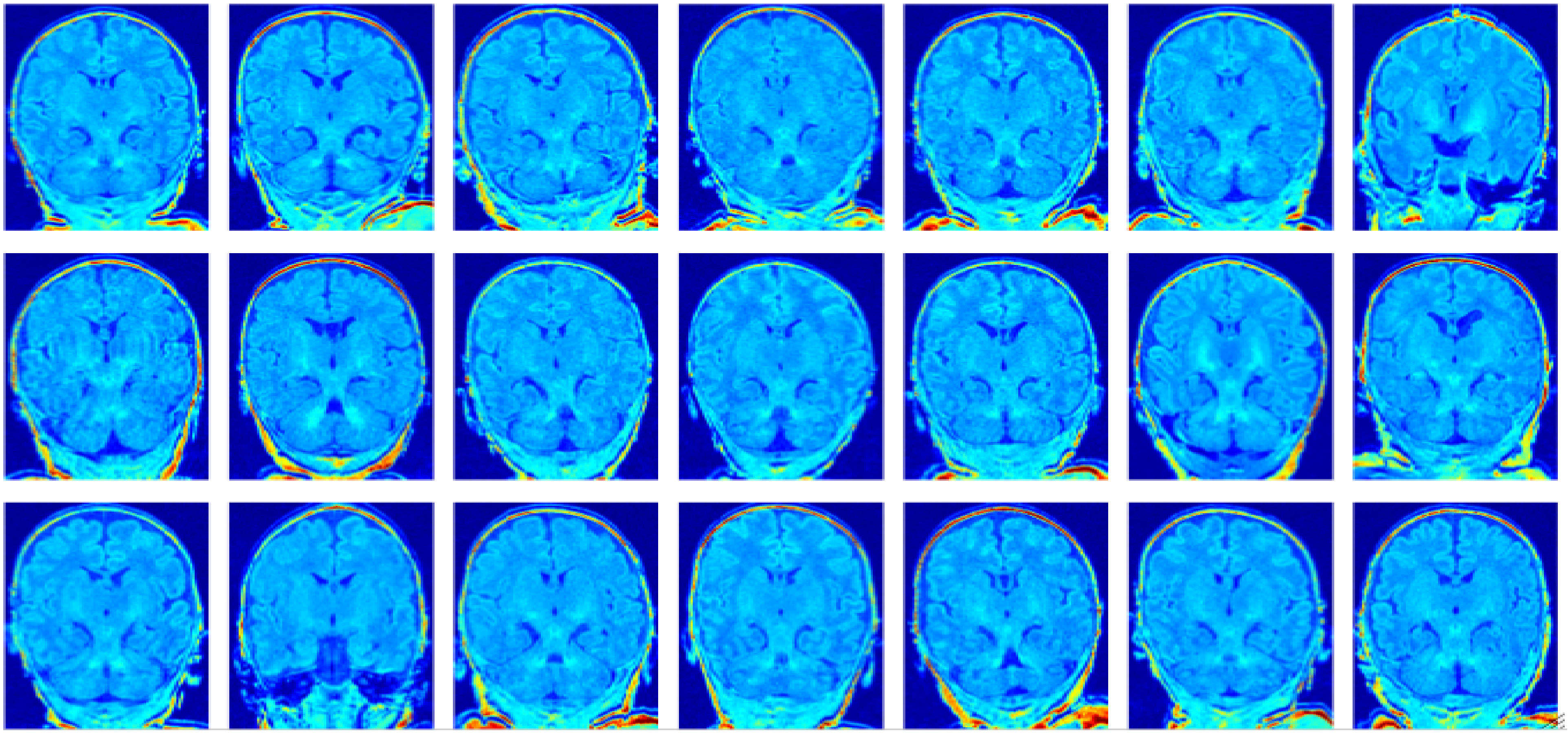
- Pixels in same location, across images, are independent.
  - When is this not true?
    - Systematic bias fields.
- Pixels in same image are independent, given their location.
  - Clearly not true, but again, doesn't seem to matter.
- Bias fields are truly bandlimited.

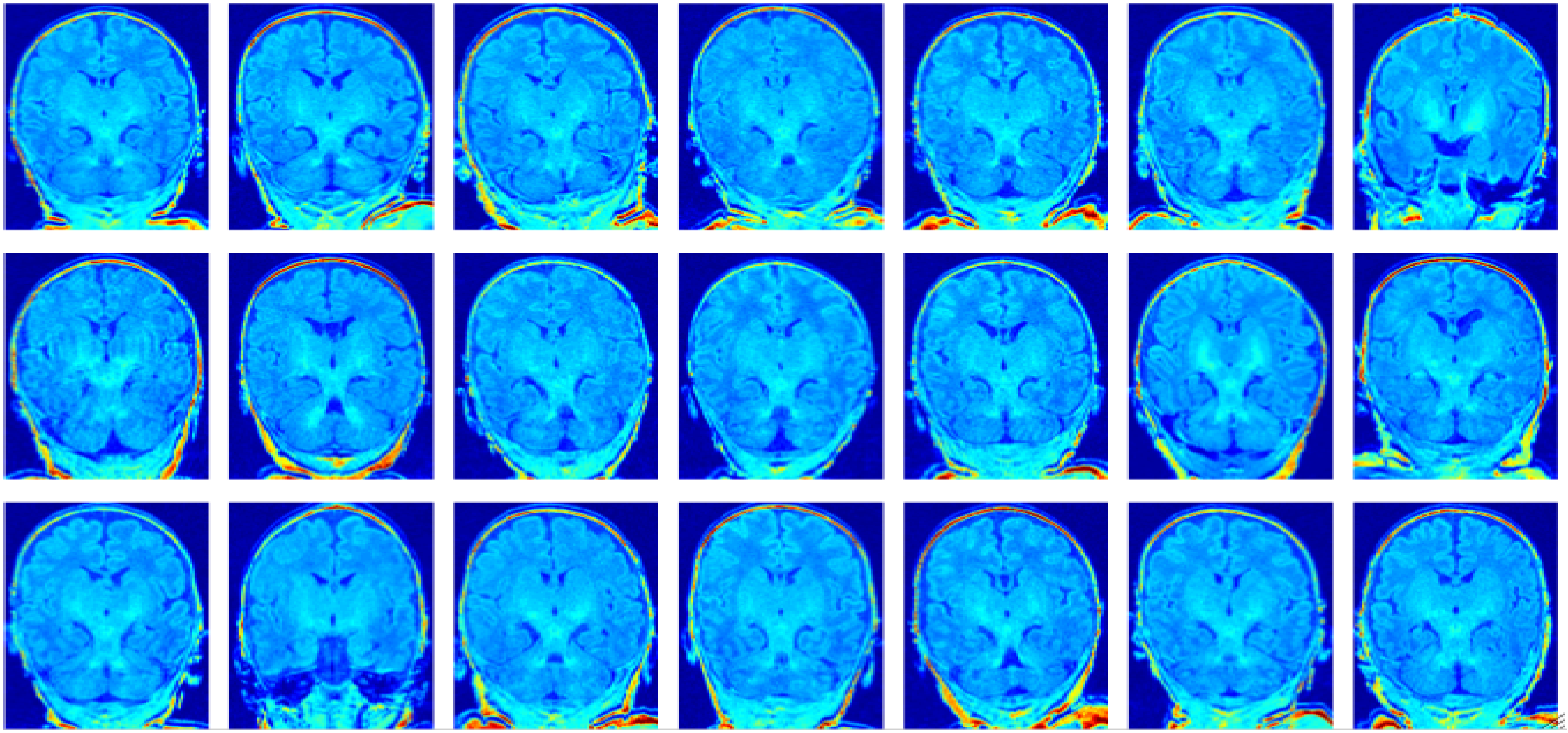
Some Other Recent Approaches

- Minimize entropy of intensity distribution in single image
  - Viola (95)
  - Warfield and Weisenfeld extensions (current)
- Wells (95)
  - Use tissue models and maximize likelihood
  - Use Expectation Maximization with unknown tissue type
- Fan (02)
  - Incorporate multiple images from different coils, but same patient.

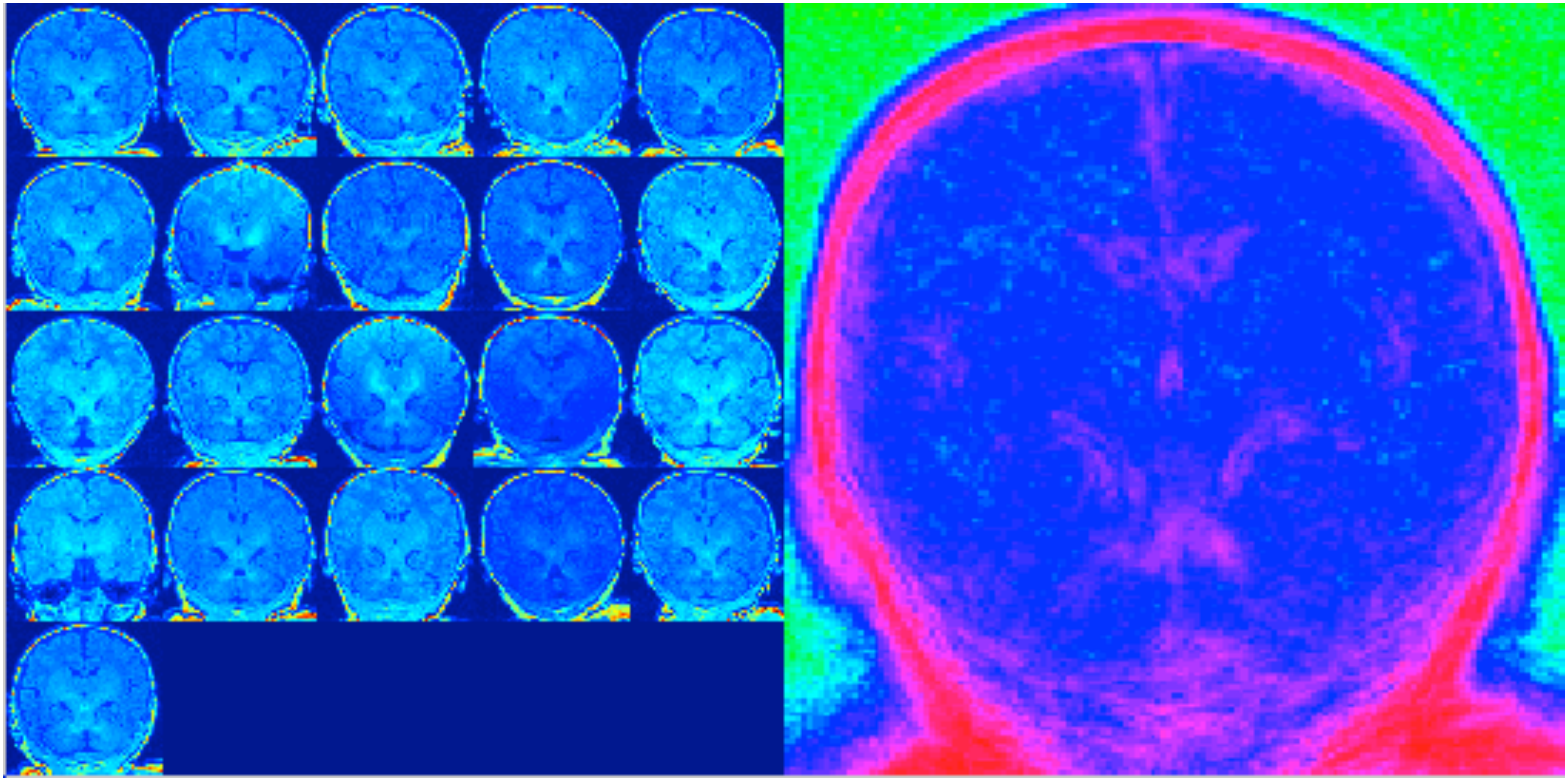
Potential difficulties with single image method

- If there is a component of the brain that looks like basis set, it will get eliminated.
- Does this occur in practice?
  - Yes!





MRI Bias Removal



Summary

- Congealing: joint alignment of images
- Learning from one example
  - Use congealing to learn about shape changes of a class
  - Transfer shape change knowledge to new classes
- Remove unwanted spatial transformations and brightness transformations from medical images
- Define notions of central tendency in a data driven manner
- Build alignment machines (funnels) that have few local minima with no labeled examples.
- Improve classification performance

UMassAmherst

Thanks!